TECHNICAL MEMORANDUM

MULTISPECTRAL SCANNER DATA PROCESSING OVER SAM HOUSTON NATIONAL FOREST: A PROGRESS REPORT

By

C. A. Reeves

E. P. Kan

APPROVED: _I. E. Duggan_
I. E. Duggan
Acting Supervisor
Forestry Applications Section

December 1974

CONTENTS

1.0 SUMMARY

This memorandum reports on the progress of the Forestry Applications Project on Timber Resources (FAP/TR), Task II.2.1 and Task II.2.2, as of December 1974 (ref. 1,2). Using existing computer processing techniques, two sets of multispectral scanner (MSS) data were used in the analysis; one set being a 3-date Earth Resources Technology Satellite (ERTS-1) MSS data, the other being a 1-date aircraft MSS data. Both data sets were over a selected area in Sam Houston National Forest (SHNF).

2.0 OBJECTIVES

Being part of FAP/TR, there are two objectives to Task II.2.1 of Timber Type-Mapping - ERTS-1 MSS, and Task II.2.2 of Timber Type-Mapping - Aircraft MSS/24:

- To evaluate existing computer processing techniques in mapping timber types using ERTS-1 and aircraft data.

- To afford the opportunity to open up new research and development (R&D) areas, while forestry data is being closely examined.

3.0 SCOPE

The "Edit 9" forest scene (ref. 2) over SHNF, which covers approximately 11 square kilometers (5 square miles), was examined. Three dates of ERTS-1 MSS data were used, 27 November 1972 (ERTS-1 scene ID #1127-16253), 25 February 1973 (ID #1217-16254), and 8 May 1973 (ID #1289-16254).

1

MSS data collected by Mission M230 flown by the C-130 aircraft on 21 March 1973 at 10,000 feet altitude was used. The Bendix 24-channel scanner was on board.

There are seven timber types/condition classes in Edit 9. These features are some of those intended for study in FAP/TR (ref. 1). Using the same notation as in ref. 1, these seven features are tabulated in Table I.

The size of ERTS-1 data for Edit 9 is approximately 30 × 80 pixels (30 lines and 80 pixels/line). The size of the aircraft MSS-24 data for the same area is approximately 250 × 700 pixels.

The spectral coverage by the ERTS-1 scanner is in Table II. The 24 channels of the Bendix MSS/24 scanner cover the visible and infrared spectrum. However, only 12 of these channels were operating during M230. They are channels 1-11 and 13 on the scanner; but they are renumbered here 1-12. The spectral coverages of these channels are in Table III.

4.0   TECHNICAL APPROACH

The raw MSS data was preprocessed before classification/ analysis. For aircraft MSS/24 data, the preprocessing comprised (a) calibration, using the BENMSS and REFORM computer programs; (b) scan-angle correction, using the LYKIT, LUVIT, LUMPIT and FIXIT computer programs; and (c) registration to ground, using the ADAS and ADJUST computer programs (ref. 1). For ERTS-1 data, the November data was registered to ground using longitude-latitude

2

TABLE I. — TIMBER TYPES/CONDITION CLASSES OF INTEREST IN EDIT 9

Type ID	Description
1.3	Shortleaf pine, immature sawtimber
2.3	Loblolly pine, seedling and sapling, adequately stocked
2.5	Loblolly pine, immature sawtimber
2.6	Loblolly pine, mature sawtimber
3.1	Laurel oak - willow oak, immature sawtimber
4.2	Sweetgum-nuttal oak - willow oak, immature sawtimber
7.2	Cutover land, not site prepared

TABLE II. – SPECTRAL COVERAGES OF THE
4 CHANNELS ON ERTS-1 MSS

Channel #	Spectral Coverage (micrometer)
1	0.5 - 0.6
2	0.6 - 0.7
3	0.7 - 0.8
4	0.8 - 1.1

TABLE III. – SPECTRAL COVERAGES OF THE 12 CHANNELS
OF THE EDIT 9 MSS AIRCRAFT DATA

Channel #	Spectral Coverage (micrometer)
1	0.375 - 0.405
2	0.40 - 0.44
3	0.466 - 0.495
4	0.53 - 0.58
5	0.588 - 0.643
6	0.65 - 0.69
7	0.72 - 0.76
8	0.770 - 0.810
9	0.82 - 0.88
10	0.981 - 1.045
11	1.20 - 1.30
12	2.10 - 2.36

4

information on the Earth Resources Interactive Processing System (ERIPS). Then the February and May data were registered and composed with the November data, on ERIPS, to form a 12-channel data set registered to ground.

Two classification/analysis approaches were then taken:

    I   Classification by training class statistics.

    II   Classification by representative cluster statistics of training class.

ERIPS was used for data processing. Rectangular and nonrectangular training fields were picked from the television display of ERIPS. Statistics of these fields and thus classes were then calculated by the system.

In approach I, the best set of channels from the available set of channels were obtained by a "divergence" calculation (ref. 3, 4) on the training class statistics. For single-data ERTS data analysis, no channel selection was made. For aircraft data analysis, the best six of the available 12 channels were chosen. Six channels were chosen to provide sufficient separation between classes. The training class statistics and the best channel set were then used in the classification of the data. The classification rule on ERIPS is a maximum likelihood classification (ref. 3, 5).

In approach II, each training field of each class was clustered into five clusters. Five was chosen arbitrarily, with the contention that all possible heterogeneity would be

described by the 5-cluster representation. The three most significant clusters of each training class were then selected manually; this was achieved by overlaying cluster statistics plots one on the top of another for identifying equivalent clusters, and then, insignificant clusters were eliminated by checking the population of the cluster in the training field/class. The cluster statistics thus chosen were then used for channel selection and then classification on ERIPS.

5.0 RESULTS

5.1 Data Processing Results

5.1.1 <u>MSS/24 Data Processing</u>. There were three main categories of activities performed on Edit 9. A 3-channel color rendition of Edit 9 is shown in Figure 1 with the ground-truth timber stand boundaries; the boundaries were extracted from U.S. Forest Service information.

<u>Preprocessing</u>

As planned in ref. 1, the original Edit 9 was calibrated, scan-angle corrected and registered to ground. The resulting data resolution, i.e. size on ground of one pixel, is $(8.47$ meters$)^2$. A 3-channel color rendition of this resulting data is shown in Figure 1. The imagery before preprocessing is not available at the time of this reporting.
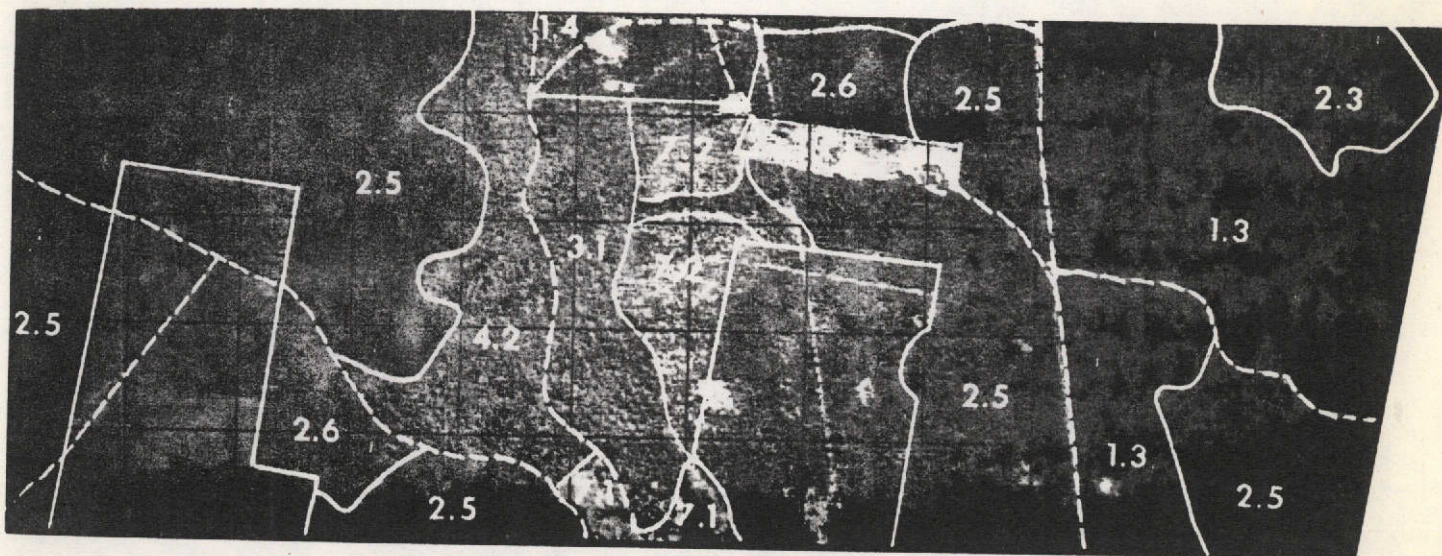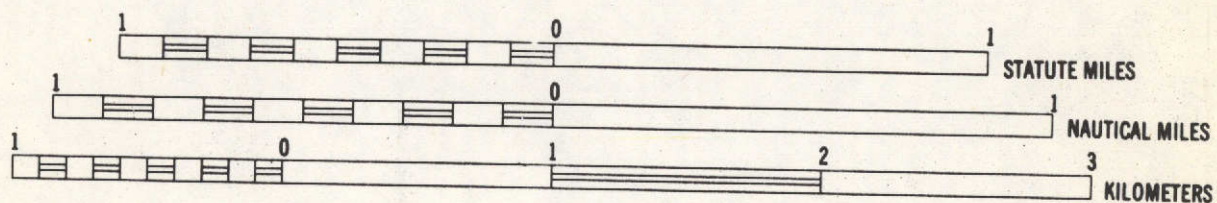
APPROXIMATE SCALE 1:28,000

Figure 1. — Timber stand and compartment map over Sam Houston National Forest Edit 9; a 3-channel color rendition of the MSS/24 data of M230. (Refer to table I for timber type codes.)

Classification Approach I

Training fields of all seven classes (c.f. section 3.0) were picked all over Edit 9. Divergence calculations produced the optimum set of channels 2, 5, 6, 7, 10 and 12. Classification of the entire Edit 9 using this channel set, class statistics and using 2 percent, 5 percent, 10 percent, and 12 percent thresholds, resulted in classification maps. Figure 2 is a generic classification map. These maps were then interpreted by a skilled photo-interpreter who delineated the boundaries of homogeneous areas, which are also shown by black lines in figure 2. Visual judgement by ten different individuals led to the conclusion that the classification map corresponding to 5 percent thresholding was the best classification map, compared to the ground-truth map of figure 1. It is also obvious from the maps that the 2.5 fields (refer to table I for timber type codes such as 2.5) on the right are considerably confused with the 1.3 fields, and are spectrally different from the 2.5 fields on the left of the scene. This suggests that the scan-angle correction previously performed on the raw Edit 9 data is inadequate.

Classification Approach II

Each training field of each class was clustered into five clusters on ERIPS using the settings of PERCENT = 0.8, SEP = 1.0, STDMAX = 0.5, DLMIN = 3.2, NMIN = 5 and ITMAX = 10 in the iterative clustering procedure. By the process stated in section 4.0, the three most significant clusters of each class were identified, and then used to classify the individual training classes. The resulting classification images were found to contain more "holes" than the maps
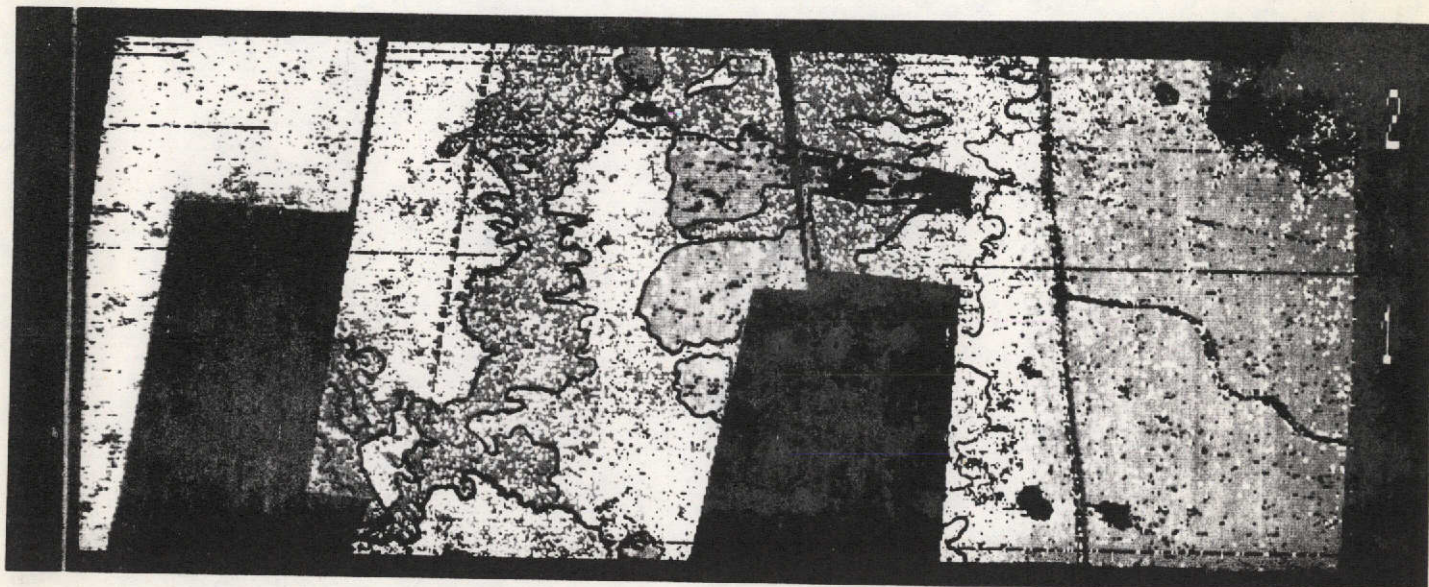
8

Figure 2. — A classification map over Edit 9. Black lines
denote timber stand boundaries delineated by a
photo-interpreter. Interpretations for colors
are: yellow = 1.3; red = 2.3; turquiose = 2.5;
blue = 2.6; white = 3.1; green = 4.2; brown = 7.2.
(Refer to table I for timber type codes.)

obtained by classifying training classes using the training class statistics; "holes" refer to thresholded pixels, i.e., pixels considered not to belong to the class. Due to technical difficulties, classification of the entire Edit 9 by cluster statistics of all seven classes has not been performed at the time of reporting.

5.1.2 <u>ERTS-1 Data Processing</u>. The three ERTS-1 data sets were preprocessed and classified on ERIPS.

Preprocessing

The November, February and May ERTS-1 data sets were registered to one another and to ground, as discussed in section 4.0. The resulting data resolution, i.e. size on ground of one pixel, is $(80 \text{ meters})^2$.

Classification

Only Approach I was taken, i.e. classification by training class statistics. Approach II of cluster-statistics-classification was not performed because it was felt that the training data did not have sufficient variability to call for clustering. Classification was done season-by-season, i.e. 1-date (4 channels) at a time. All three season classifications showed that pine and hardwood could be separated from one another; pine pertains to the totality of 1.3, 2.3, 2.5, 2.6, and hardwood pertains to 3.1 and 4.2. The May classification seemed to be better than the November and February classification. And, the 2.3 and 7.2 class were very well distinguished from each other and from the other features in all seasons. Hardcopies of classification maps were not available at the time of this reporting.

10

5.2   New Research and Development Areas.

As a result of the intensive data processing activities on forestry data, problem areas that require research and development opened up and became more well defined. Four such problem areas are listed as follows:

(a)   Mixture modeling — The extreme complexities in the aircraft data (at $(8 \text{ meters})^2$ resolution) call for new and improved modeling of forestry data, especially at high resolution. Also, a systematic, algorithmic way to identify equivalent clusters, significant and insignificant clusters need to be developed. These problems are being investigated and partial solutions are reported in ref. 6.

(b)   Data resolution versus classification — It appears that classification is made difficult by the high resolution in the aircraft data. It also appears that classification is gross in the low resolution ERTS-1 data. A natural question arises: "What is the optimal data resolution for best forestry analysis?" Modeling and simulation of data with intermediate resolution are being performed. Some results are reported in ref. 7.

(c)   Boundary finding — Boundaries between timber stands in classification maps normally are fuzzy. Boundary finding is an art in human photo-interpretation, and is far from being well defined in machine processing. Ideas are being explored in using lower resolution data for boundary finding.

(d) Scan-angle correction — The Edit 9 aircraft data is found to be inadequately corrected for scan-angle effects. Signatures on the left side of the scene does not seem to be extendable to the right side. Closer look at the process of scan-angle correction is called for.

6.0 CONCLUSION

The Edit 9 forest scene over Sam Houston National Forest has been examined intensively using aircraft and ERTS-1 data, for the two objectives of FAP/TR: (a) evaluate existing computer techniques for forestry data processing; (b) afford an opportunity to open up new research and development areas.

Although the first objective is not yet completely fulfilled, the 6-month's study in this reporting period has opened up at least two major research and development areas. These are (a) modeling of forestry data, and (b) data resolution versus classification study; these works have been reported in ref. 6 and 7.

Should manpower be made available in the next 6 months to these two Tasks II.2.1 and II.2.2 of FAP/TR, data processing and analysis effort should be expended to complete what has been left undone here, and also to study other available data sets of Sam Houston National Forest.

7.0 REFERENCES

1. Kan, E. P.: The Forestry Applications Project on Timber Resources (FAP/TR): The FY75 Investigation Plan. Lockheed Electronics Co., Inc., Houston, Texas, Tech. Rep. #LEC 3567, June 1974.

2. Kan, E. P.: The Forestry Applications Project on Timber Resources (FAP/TR): The FY75 Interim Report. Lockheed Electronics Co., Inc., Houston, Texas, Tech. Rep. #LEC 5135, Dec. 1974.

3. Earth Resources Interactive Processing System: A User's Manual. International Business Machines Corporation, Federal Systems Division, Houston, Texas, Contract No. NAS 9-996, Dec. 1972.

4. Marill, T. and Green, D. M.: On the Effectiveness of Receptors in Recognition Systems. IEEE Transactions on Information Theory, Vol. IT-9, Jan. 1963.

5. Anderson, T. W.: An Introduction to Multivariate Statistical Analysis. John Wiley and Sons, New York, N.Y. 1958.

6. Basu, J. P. and Kan, E.P.: Statistical Modeling and Analysis of Forestry Data: Pure and Mixed Pixels. Part I. Lockheed Electronics Co., Inc., Houston, Texas, Tech. Memo. #LEC 5185, Dec. 1974

7. Kan, E. P. and Ball, D. L.: Data Resolution versus Forestry Classification. Lockheed Electronics Co., Inc., Houston, Texas, Tech, Rep. #LEC 5168, Dec. 1974.